PET Image Reconstruction Using Simulated Annealing

Erik Sundermann*

Department of Electronics and Information Systems,
University of Ghent, St. Pietersnieuwstraat 41, B-9000 Ghent, Belgium

Ignace Lemahieu†

Department of Electronics and Information Systems,
University of Ghent, St. Pietersnieuwstraat 41, B-9000 Ghent, Belgium

ABSTRACT

In Positron Emission Tomography (PET) images have to be reconstructed from noisy projection data. The noise on the PET data can be modeled by a Poisson distribution. The development of statistical (iterative) reconstruction techniques addresses the problem of noise.

In this paper we present the results of introducing the simulated annealing technique as a statistical reconstruction algorithm for PET. We have succesfully implemented a reconstruction algorithm based upon simulated annealing, with paying particular attention to the fine-tuning of various parameters (cooling schedule, granularity, stopping rule, ...). In addition, we have developped a cost function more appropriate to the noise statistics (e. g. Poisson) and the reconstruction method (e. g. ML). The comparison with other reconstruction methods using computer phantom studies proves the potential power of the simulated annealing technique for the reconstruction of PET-images.

1 INTRODUCTION

Positron Emission Tomography is a tomographic method to display metabolic activity in a slice through a patient's body. The particular construction of the PET scanner and the use of a radioactive tracer entail the modeling of the data by a Poisson distribution.

The reconstruction method most commonly used today in PET is the Filtered Backprojection (FB) algorithm.[1] This reconstruction technique is based on a Fourier Transform algorithm and is extremely fast. However, since FB is a deterministic algorithm and hence does not account for statistical fluctuations in the measurements, the obtained reconstructed images can suffer from very annoying stripe-like artifacts.

* Research supported by a grant from the IWT, Brussels, Belgium
† Senior research associate with the NFWO, Brussels, Belgium

0-8194-1782-3/95/$6.00

To address the problem of noise, the study of statistical (iterative) reconstruction techniques has received much attention in the past few years. One of the most popular methods is the Maximum Likelihood Expectation Maximization (ML-EM) algorithm,[2,3] which searches the image that maximizes the likelihood of the data. However, when the algorithm is iterated too long, the reconstructed image starts to degrade.[4]

Maximization algorithms (e. g. conjugate gradient, EM) achieve convergence by monotonically increasing some metric (e. g. a likelihood) at every iteration. Simulated annealing is a Monte Carlo technique which allows occasional negative increments of the cost function so that it can avoid getting trapped in local minima. The technique has been proposed earlier by Kearfott *et. al.* [5] and Webb[6] for SPECT, but it has never been applied to PET before.

2 SIMULATED ANNEALING

Simulated annealing was introduced by Metropolis *et. al.* [7] and is used to approximate the solution of very large combinatorial optimization problems[8] (e. g. NP-hard problems). The technique originates from the theory of statistical mechanics and is based upon the analogy between the annealing of solids and solving optimization problems.

Let us assume we are looking for the configuration that minimizes a certain cost function E. The algorithm can then be formulated as follows.[9] Starting off at an initial configuration, a sequence of iterations is generated. Each iteration consists of the random selection of a configuration from the neighbourhood of the current configuration and the calculation of the corresponding change in cost function ΔE. The neighbourhood is defined by the choice of a generation mechanism, i. e. a "prescription" to generate a transition from one configuration into another by a small perturbation. If the change in cost function is negative, the transition is unconditionally accepted; if the cost function increases the transition is accepted with a probability based upon the Boltzmann distribution

$$\mathrm{P}_{acc}(\Delta E) \sim \exp(-\frac{\Delta E}{kT}),$$

where k is a constant and the temperature T is a control parameter. This temperature is gradually lowered throughout the algorithm from a sufficiently high starting value (i. e. a temperature where almost every proposed transition, both positive and negative, is accepted) to a "freezing" temperature, where no further changes occur. In practise, the temperature is decreased in stages, and at each stage the temperature is kept constant until thermal quasi-equilibrium is reached. The whole of parameters determining the temperature decrement (initial temperature, stop criterion, temperature decrement between successive stages, number of transitions for each temperature value) is called the cooling schedule.

Consequently the four key "ingredients" for the implementation of simulated annealing are:

- the definition of configurations;
- a generation mechanism, i. e. the definition of a neighbourhood on the configuration space;
- the choice of a cost-function;
- a cooling schedule.

We will now discuss the particularities of the annealing algorithm in the case of reconstruction of PET images.

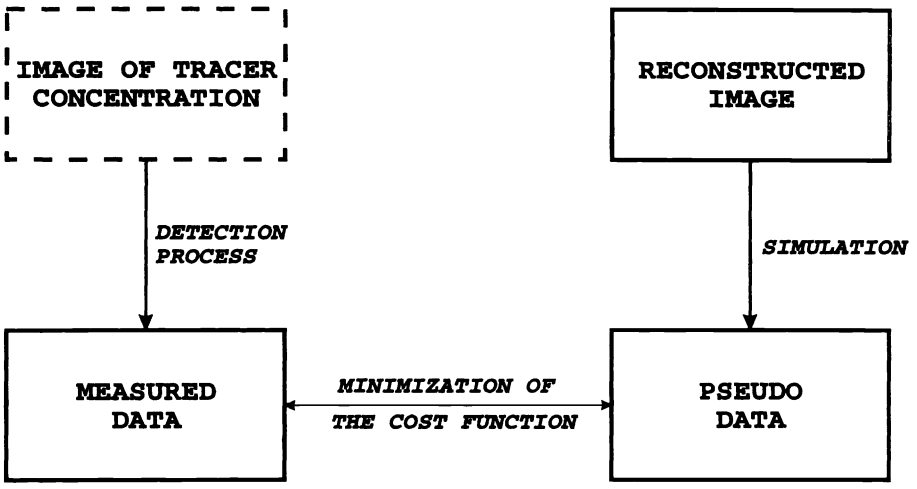

Figure 1: Reconstruction scheme for simulated annealing.

3 A RECONSTRUCTION ALGORITHM BASED ON SIMULATED ANNEALING

3.1 Principle

The simulated annealing reconstruction algorithm for PET can be formulated as follows. We will iteratively reconstruct an image that fits best the measured data $\mathbf{P}^m$. To do so, we will calculate at each iteration step the pseudo data $\mathbf{P}^p$ that correspond to the present state of the reconstructed image. We assume that, by minimizing the difference between the measured data and the pseudo data, the reconstructed image will converge towards the sought-after original image. Therefore we choose as cost function a function that expresses the difference between both data sets. This reconstruction scheme is shown in Fig. 1.

At each iteration step, the intensities of one or a few pixels are altered. We calculate the corresponding change in cost function ΔE, and decide upon this difference in cost function whether the proposed transition is accepted. The "crystallization" of the reconstructed image throughout the algorithm can be seen in Fig. 2; the evolution of the cost function is shown in Fig. 3. Since we use software phantoms, we can also display the error function of the reconstructed image, which is the least squares distance between the original and the reconstructed image.

3.2 Configurations

We have performed studies on software generated phantom images. The pixel intensities are integer values ranging between 0 and 10000. To reduce reconstruction times, we have experimented with relatively small 64 x 64 images, since there is no reason why the obtained results should not be valid for larger dimensions. We have furthermore restricted our experiments to the reconstruction of noisefree data. To be able to evaluate different properties of the reconstructed images (such as contrast, smoothness) we have used a number of different phantoms.

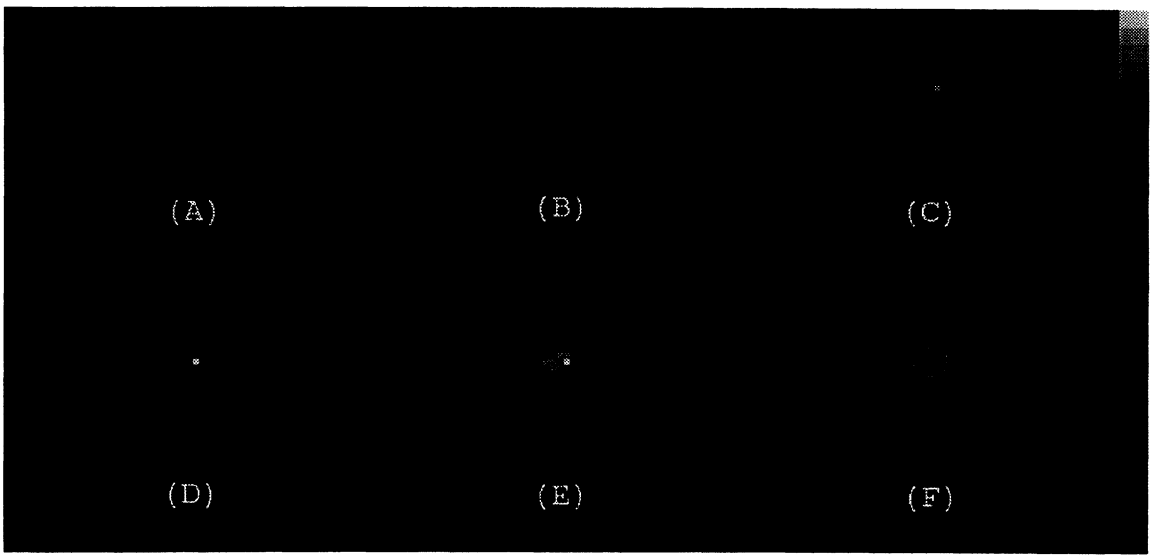

Figure 2: Reconstructed image after (A) 500.000, (B) 1 million, (C) 1.5 million, (D) 2 million, (E) 3 million and (F) 5 million iterations.

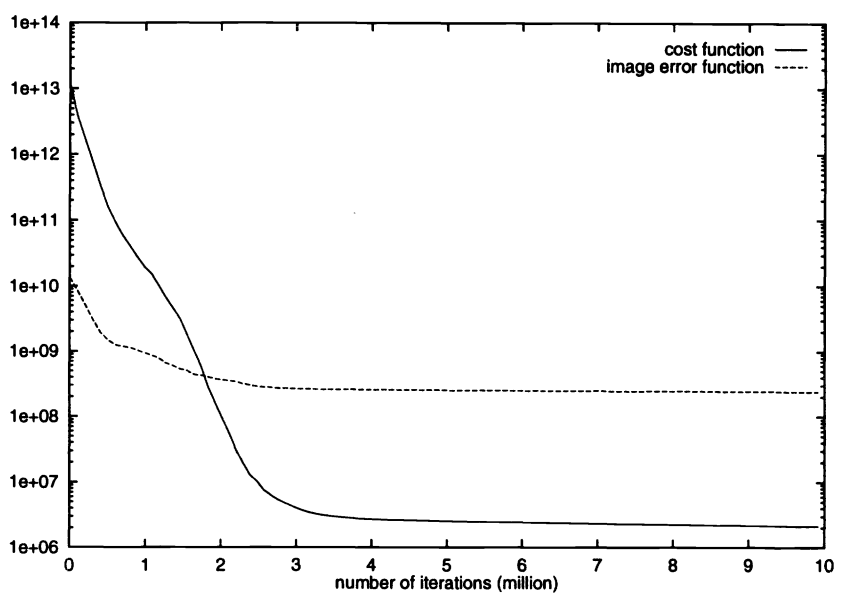

Figure 3: Evolution of the cost function and the error function of the reconstructed image.

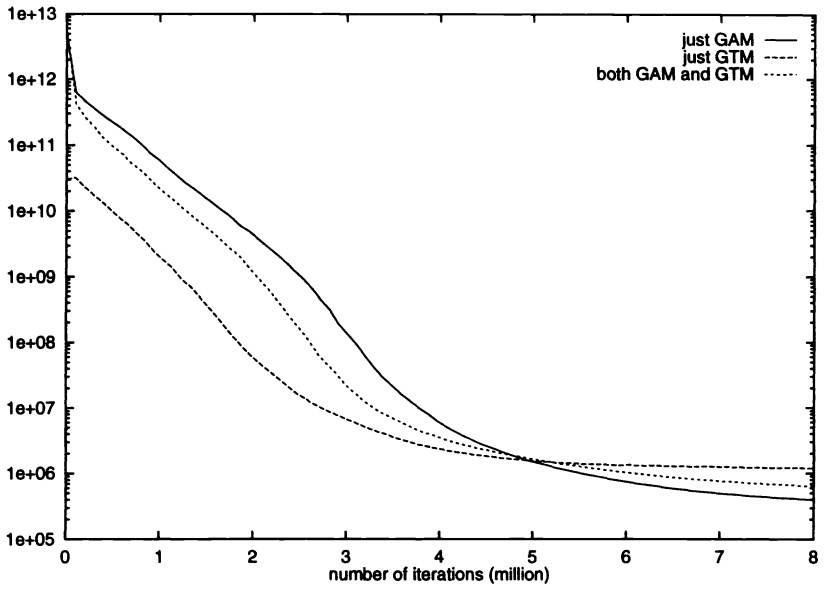

Figure 4: Evolution of the cost function for different pixel updating methods; the GTM reconstruction started from an initial image with correct total intensity.

3.3 Generation mechanism

As explained above, the generation mechanism controls the transition from one image to a neighbouring one. This is achieved by adjusting the intensities of one or more randomly chosen pixels at each iteration step. We therefore have to decide on the number of pixels to be changed simultaneously, the choice of pixels and the amount of intensity to be changed (i.e. the grainsize).

Pixel updating We considered two possible approaches to the first problem, based upon Frieden *et. al.* [11] There is the Grain Allocation Method (GAM), which allocates a grain of intensity to one pixel, and there is the Grain Transfer Method (GTM), which transfers a grain from one pixel to another. Since GTM conserves the total image intensity (which is initially unknown), we constructed a third method which uses both GAM and GTM at the same time. Our experiments showed that GAM performs marginally better than the other two methods, as can be seen in Fig. 4.

Pixel choice PET measurements have to be corrected for attenuation in the scanned object. This attenuation correction is usually derived from an extra transmission scan. Huang *et. al.* [12] describe a technique to replace the extra transmission scan by a calculated attenuation correction, based upon the contour of the scanned object. Since each zero-element in the measured data implies a strip of empty pixels in the image at the corresponding angle, the contour can easily be found.

We have used this technique to exclude whole regions of the image from the generation mechanism, thus improving the efficiency of each iteration step. The resulting evolution of the cost function is displayed in Fig. 5. As can be expected, this technique only improves the efficiency of the generation mechanism, but not the quality of the reconstructed image.

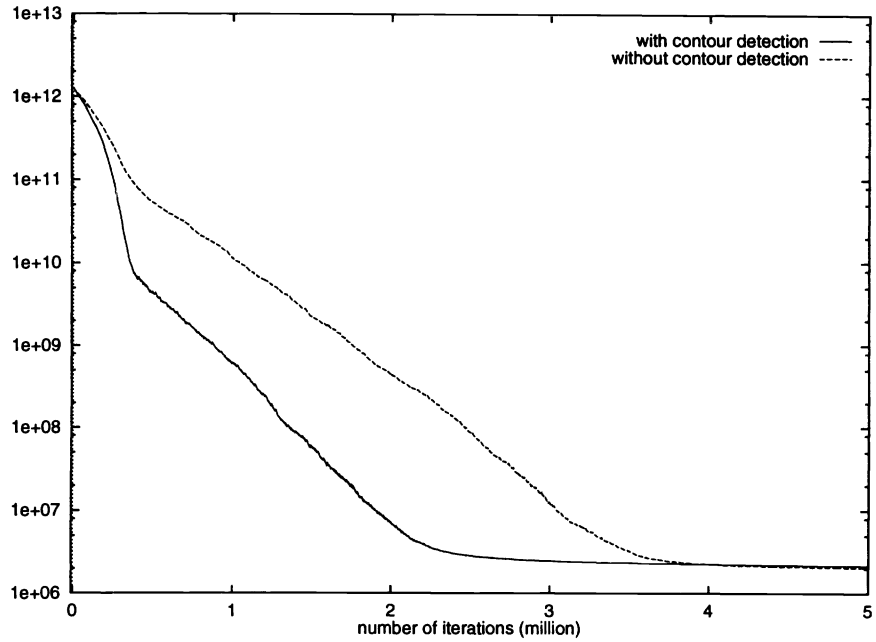

Figure 5: Evolution of the cost function with and without preliminary contour detection.

Grainsize It is clear that the algorithm must allow both positive and negative grainsizes, but must forbid negative image intensities. Most authors suggest choosing grainsizes from a uniform distribution in the interval $[-\beta, +\beta]$. As can be seen in Fig. 6, our experiments show that better results can be obtained by not keeping the boundary parameter β constant, but letting the algorithm itself decide upon the value for β throughout the process. This can be achieved by keeping track of the mean accepted grainsize, and choosing the boundary parameter accordingly.

3.4 Cost function

We first performed experiments using the classical cost function

$$E = \sum_{i,j} (P_{ij}^m - P_{ij}^p)^2, \qquad \forall i, j = 1, \ldots, 64,$$

which is the mean squared distance between $\mathbf{P}^m$ and $\mathbf{P}^p$. This approach is suggested by most authors[5,6] and produces good results for a number of different testphantoms. The reconstructed images show that simulated annealing, in comparison to other reconstruction algorithms, is very good at reconstructing edges between image-areas with a constant intensity, but has more difficulties in reconstructing these constant areas themselves. This problem can be addressed by adding an appropriate smoothness factor to the cost function.[10]

We later derived a new cost function, based upon the maximum likelihood principle. Maximization of the likelihood $\mathrm{P}(\mathbf{I}|\mathbf{P}^m)$, with $\mathbf{I}$ the reconstructed image, implies maximization of $\prod_{i,j} \mathrm{P}(P_{ij}^m|\mathbf{I})$. Taking into account the Poisson-character of the detection process, we find that

$$\mathrm{P}(P_{ij}^m|\mathbf{I}) = e^{-P_{ij}^p} \frac{(P_{ij}^p)^{P_{ij}^m}}{P_{ij}^m!}.$$

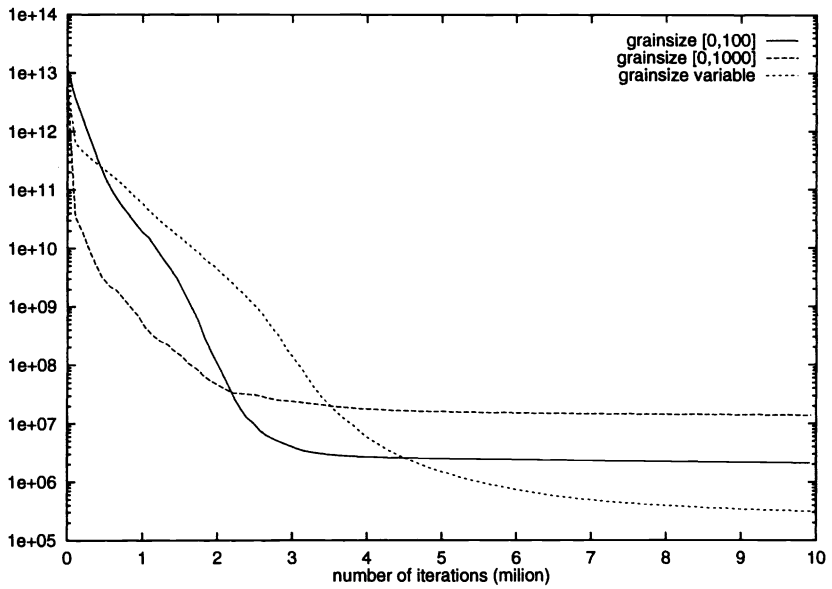

Figure 6: Evolution of the cost function for different grainsize settings.

After taking the logarithm of the likelihood and omitting the constant term, the cost function reduces to

$$E = \sum_{i,j}(P_{ij}^m \ln P_{ij}^p - P_{ij}^p) \qquad \forall i,j = 1,\ldots,64.$$

This new cost function achieves reconstructed images of the same quality as the classical cost function, but clearly shows faster convergence in the early stages of the reconstruction. Since this new cost function is based upon the statistical character of the detection process, we expect it to outperform the classical cost function when applied to noisy data.

3.5 Cooling schedule

Starting temperature The initial temperature should be chosen so that about 80% of all positive transitions (i.e. transitions which increase the cost function) are accepted.[9] This temperature is strongly dependent on the measured data and the initial image. Most authors suggest a trial and error method for determining this starting temperature. This can be done for example by running the algorithm shortly for one temperature value, calculating the corresponding acceptance rate and adjusting the temperature value accordingly until a rate of about 80% is achieved.

We have developed a criterion which calculates the starting temperature from only one trial run. This trial run results in a series of N positive transitions ΔE_i. We are looking for a value for kT for which

$$\sum_{i=1}^{N} \frac{1}{N} \exp(-\frac{\Delta E_i}{kT}) = 0.8.$$

After neglecting the higher order terms of the Taylor-expansion of the exponential, we find that

$$kT = 5(\Delta E_i)_{mean}.$$

Experiments show that this value leads to acceptance rates of about 82%, which means that this method is a very good estimator for the starting temperature.

Temperature decrement between successive stages The temperature is decreased by multiplication with a factor α. There is actually a trade-off between temperature decrement between stages and the number of iterations per stage. For example when the temperature is kept constant long enough for the equilibrium to be reached, greater decrements are allowed. Best values for α are between 0,8 and 0,98.[9]

Number of iterations per stage Frequently used criteria are a constant number of iterations, or iterating until a constant number of transitions is accepted. Experiments show that better results are achieved by considering the physical background of simulated annealing and the concept of thermal quasi-equilibrium. This means keeping the temperature constant until the cost function has reached a constant value (or is oscillating around this constant value).

Stop criterion Since the emphasis of our research was on the quality of the reconstructed images, we have not devoted much attention to the development of a workable stop criterion. It is clear however that when the equilibrium values of the cost function for successive stages are constant themselves, the iteration process can be stopped.

4 RESULTS AND CONCLUSION

We have found the simulated annealing algorithm to reconstruct images of good quality for noiseless data. As already mentioned, it is notable that the reconstructed images show a high contrast in comparison to other reconstruction techniques (Fig. 7). These results certainly justify the continuation of our research for the reconstruction of noisy images. The main problem with simulated annealing is the large reconstruction time. Since it has been proven that simulated annealing is extremely suited for parallellization,[13] we will try to reduce the reconstruction time by parallellizing the algorithm.

5 REFERENCES

[1] A.C. Kak, M. Slaney, *Principles of Computerized Tomographic Imaging*, New York : IEEE Press, 1988.

[2] L.A. Shepp, Y. Vardi, *Maximum likelihood reconstruction for emission tomography*, IEEE Trans. Med. Imaging, Vol. 1, pp. 113-122, 1982.

[3] P. Desmedt, K. Thielemans, I. Lemahieu, F. Vermeulen, D. Vogelaers, F. Colardyn, *Measured attenuation correction using the Maximum Likelihood Algorithm*, Medical Progress through Technology, Vol. 17, pp. 199-204, 1991.

[4] D.L. Snyder, M.I. Miller, L.J. Thomas, D.G. Politte, *Noise and edge artifacts in ML reconstructions for emission tomography*, IEEE Trans. Med. Imaging, Vol. 6, pp. 228-238, 1987.

[5] K.J. Kearfott, S.E. Hill, *Simulated annealing image reconstruction method for a pinhole apperture single photon emission computed tomograph (SPECT)*, IEEE Trans. Med. Imaging, Vol. 9, pp. 128-143, 1990.

[6] S. Webb, *SPECT Reconstruction by Simulated Annealing*, Phys. Med. Biol., Vol. 34, No. 3, pp. 259-281, 1989.

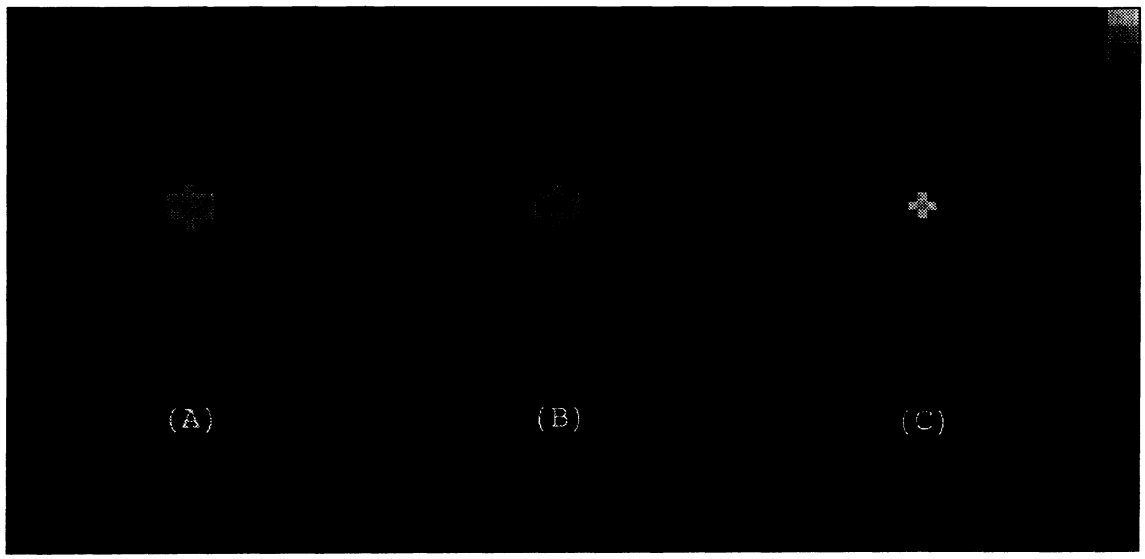

Figure 7: Comparison between (A) the original image and the reconstructed images with (B) simulated annealing and (C) maximum likelihood.

[7] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, *Equation of State Calculations by Fast Computing Machines*, J. of Chem. Phys., Vol. 21, No. 6, pp. 1087-1092, 1953.

[8] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, *Optimization by simulated annealing*, Science 220, pp. 671-680, 1983.

[9] P.J.M. van Laarhoven, E.H.L. Aarts, *Simulated Annealing: Theory and Applications*, Kluwer Ac. Publ., Dordrecht, 1987.

[10] S. Geman, D. Geman, *Stochastic relaxation, Gibbs distributions and Bayesian restoration of images*, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 6, pp. 721-741, 1984.

[11] B.R. Frieden, *Restoration of Pictures by Monte-Carlo Allocation of Pseudograins*, Proc. 2nd Int. Joint Conf. on Pattern Recognition (IEEE, New York), Copenhagen, pp. 141-142, 1974.

[12] S.C. Huang, R.E. Carson, M.E. Phelps, E.J. Hoffman, H.R. Schelbert, D.E. Kuhl, *A Boundary Method for Attenuation Correction in Positron Computed Tomography*, J. Nucl. Med., Vol. 22, No. 7, pp. 627-637, 1981.

[13] K.A. Girodias, H.H. Barrett, R.L. Shoemaker, *Parallel simulated annealing for emission tomography*, Phys. Med. Biol., Vol. 36, No. 7, pp. 921-938, 1991.